How to Learn Accurate Grid Maps with a Humanoid

Cyrill Stachniss   Maren Bennewitz   Giorgio Grisetti   Sven Behnke   Wolfram Burgard

Abstract— **Humanoids have recently become a popular research platform in the robotics community. Such robots offer various fields for new applications. However, they have several drawbacks compared to wheeled vehicles such as stability problems, limited payload capabilities, violation of the flat world assumption, and they typically provide only very rough odometry information, if at all. In this paper, we investigate the problem of learning accurate grid maps with humanoid robots. We present techniques to deal with some of the above-mentioned difficulties. We describe how an existing approach to the simultaneous localization and mapping (SLAM) problem can be adapted to robustly learn accurate maps with a humanoid equipped with a laser range finder. We present an experiment in which our mapping system builds a highly accurate map with a size of around 20 m by 20 m using data acquired with a humanoid in our office environment containing two loops. The resulting maps have a similar accuracy as maps built with a wheeled robot.**

I. INTRODUCTION

In the last few years, humanoid robots have become a popular research tool. They are assumed to offer new perspectives compared to wheeled vehicles since they are, for example, able to access different types of terrain and climb stairs. Generally, their human-like body plan helps when acting in a world designed for humans. The drawback of humanoids is that several tasks that can be easily carried out with wheeled robots are hard to achieve with legged systems. This includes, for example, stable motion with payload and the accurate execution of motion commands.

Maps of the environment are needed for a wide range of robotic applications including search and rescue, automated vacuum cleaning, home assistance, and several other service robotic tasks. Learning maps has therefore been a major research topic in the robotics community over the last decades. In the literature, the mobile robot mapping problem is often referred to as the simultaneous localization and mapping (SLAM) problem. It is considered to be a complex problem, because for localization a robot needs a consistent map and for acquiring a map a robot requires a good estimate of its location. This mutual dependency between the pose and the map estimates makes the SLAM problem hard and requires searching for a solution in a high-dimensional space. Several techniques to the SLAM problem have been developed for wheeled robots but only a few of them have been shown to work on humanoid robots.

The central question in this context is what makes the data acquired with a humanoid different from data obtained with a wheeled platform. First, wheeled platforms are typically

All authors are with the University of Freiburg, Department of Computer Science, D-79110 Freiburg, Germany

Fig. 1. A learned map by our approach using noisy and short-range laser data acquired with the humanoid Robotinho.

equipped with sensors that provide rather accurate odometry information. This yields good estimates of the relative movement or at least a reasonable starting point for local pose correction methods such as scan-matching. Compared to that, most humanoid robots (at least the ones that are affordable) do not have any odometry sensor. Furthermore, wheeled robots provide significantly more stable and smooth motion behaviors. This allows most robots to make the 2D plane assumption which means that the robot moves on a plane and the sensor is located parallel to that plane. Typically, this is not the case with a humanoid robot since they need to keep their balance at all time, even while standing. The attitude (roll and pitch angle) of the robot's sensors can easily change up to $20°$. Due to the very limited payload of most humanoid robots and to keep the motion behavior stable, sensors have to be light-weight and small. A SICK LMS sensor for example, cannot be mounted on most humanoids. Therefore, one typically has to deal with rather noisy and short-range sensor data resulting from light-weight laser scanners such as Hokuyo URGs.

In this paper, we investigate humanoid-specific adaptations of a mapping approach that has been successfully used on wheeled vehicles equipped with a SICK laser range finders. We present a variant of our Rao-Blackwellized particle filter for learning grid maps [6] that can be used on a humanoid. This includes corrections for changing attitude (roll and pitch) of the sensor, dealing with missing odometry information, and scan-matching with a few distinct features only. In contrast to other mapping system that operate on humanoid robots, our approach is able to learn maps of comparably large indoor environments. We present an experiment in which our system built a map of an environment with a size of 20 m by 20 m containing two loops. The map has a quality comparable to the ones generated by a wheeled robot. The map shown in Fig. 1 illustrates the result of our mapping system given data acquired with our humanoid robot.

II. RELATED WORK

The majority of approaches described in the SLAM literature addressed the problem of building maps with wheeled platforms. So far, only few researchers have addressed the problem of learning maps with a humanoid robot. Gutmann *et al.* [7] presented an approach to learn occupancy grid maps including elevation information with the Sony humanoid QRIO using stereo vision. In this context, they consider mapping mainly as a local problem to support collision avoidance or path planning tasks but they do not address issues such as loop-closing or place-revisiting.

A system that performs real-time localization and mapping with a humanoid robot was developed by Ozawa *et al.* [14]. Their approach is mainly based on 3D visual odometry and uses dense feature maps to estimate the position of the camera. A well-known drawback of this incremental approach is the drift created by the accumulation of errors.

There exist systems that concentrate on localization with a humanoid. Bennewitz *et al.* [1] presented an approach to visual localization of a humanoid that relies on a particle filter. Thompson *et al.* [16] performed localization with a humanoid equipped with a Hokuyo URG laser scanner. They use a known 2.5-dimensional map for a relatively small operational range of the robot. They do not suggest how to automatically learn such a map.

Solutions to the SLAM problem for wheeled vehicles often use EKFs. The effectiveness of the EKF approaches comes from the fact that they estimate a fully correlated posterior over landmark maps and robot poses [15]. Their weakness lies in the strong assumptions that have to be made on both, the robot motion model and the sensor noise. If these assumptions are violated, the filter is likely to diverge [9]. The unscented Kalman filter described in [9] is one way of better dealing with the non-linearities in the motion model of the vehicle. Moreover, the landmarks are assumed to be uniquely identifiable, even so, there exist techniques to deal with unknown data association in the SLAM context [13]. Lidoris *et al.* [10] presented an approach for motion planning in the context of EKF-based map learning with humanoids. They select gaze actions based on the expected entropy reduction in their model. They showed in simulation that such gaze actions can improve the pose estimate.

A full vision-based SLAM system that considers all 6 DoF and enables a humanoid robot to learn landmark maps has recently been presented by Davison *et al.* [2]. They extract features from a monocular camera and create a sparse map of high-quality stable features. The location of the features are tracked by applying an EKF.

Thrun *et al.* [17] describe a mapping approach that has been proven to work without odometry information. However, it requires an accurate laser range finder such as a SICK LMS sensor which cannot be carried by most humanoid robots. Furthermore, the sensor is assumed to have constant attitude angles. For robots equipped with a stereo camera, Elinas *et al.* [5] presented a SLAM system that does not need any odometry information.

In a work by Murphy [12], Rao-Blackwellized particle filters (RBPF) have been introduced as an effective means to solve the SLAM problem. Each particle in a RBPF represents a possible robot trajectory and a map. The framework has been subsequently extended by Montemerlo et al. [11] for approaching the SLAM problem with landmark maps. To learn accurate grid maps, RBPFs have been used by Eliazar and Parr [4] and Hähnel *et al.* [8]. Whereas the first work describes an efficient map representation, the second presents an improved motion model that reduces the number of required particles.

In this paper, we apply a variant of our mapping approach [6] that applies a Rao-Blackwellized particle filter with an informed proposal distribution to efficiently sample the next generation of particles. We adapted our approach to explicitly address the problems that appear in the context of humanoid robots. This includes missing odometry information, comparably noisy data from light-weight proximity sensors, as well as a non-constant attitude (roll and pitch angle) resulting from the walking behavior.

III. LEARNING MAPS WITH RAO-BLACKWELLIZED PARTICLE FILTERS

Mapping with Rao-Blackwellized particle filters has been first introduced by Murphy [12]. The goal is to estimate the trajectory of the robot as well as a map of the environment up to time t. The key idea of a Rao-Blackwellized particle filter for map learning is to separate the estimate of the trajectory $x_{1:t}$ of the robot from the map m of the environment. This is done by the following factorization

$$p(x_{1:t}, m \mid z_{1:t}, u_{1:t-1}) = \\ p(x_{1:t} \mid z_{1:t}, u_{1:t-1}) \cdot p(m \mid x_{1:t}, z_{1:t}), \quad (1)$$

where $z_{1:t}$ is the observation sequence and $u_{1:t-1}$ the odometry information. In practice, the first term of Eq. (1) is estimated using a particle filter and the second term turns into "mapping with known poses".

A particle filter requires three sequential steps to update its estimate. Firstly, one draws the next generation of samples from the so-called proposal distribution π. Secondly, one assigns a weight to each sample. The weights account for the fact that the proposal distribution is in general not equal to the target distribution. The third step is the resampling step in which the target distribution is obtained from the weighted proposal by drawing particles according to their weight.

One of the main challenges in particle filtering is to choose an appropriate proposal distribution. The closer the proposal is to the true target distribution, the more precise is the estimate represented by the sample set. Typically, one requires the proposal π to fulfill the assumption

$$\pi(x_{1:t} \mid z_{1:t}, u_{1:t-1}) = \pi(x_t \mid x_{1:t-1}, z_{1:t}, u_{1:t-1}) \\ \cdot \pi(x_{1:t-1} \mid z_{1:t-1}, u_{1:t-2}). \quad (2)$$

According to Doucet [3], the distribution

$$p(x_t^{(i)} \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, z_t, u_{t-1}) =$$

$$\frac{p(z_t \mid m_{t-1}^{(i)}, x_t^{(i)}) p(x_t^{(i)} \mid x_{t-1}^{(i)}, u_{t-1})}{p(z_t \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, u_{t-1})} \quad (3)$$

is the optimal proposal for particle i with respect to the *variance of the particle weights* that satisfies Eq. (2). This proposal minimizes the degeneracy of the algorithm (Proposition 4 in [3]). In Eq. (2), $z_{1:t-1}$ and $x_{1:t-1}$ represent the map m_{t-1}. Based on the importance sampling principle, the weights have to be computed as follows [6]

$$w_t^{(i)} = \frac{\text{target distribution}}{\text{proposal distribution}} \quad (4)$$

$$= w_{t-1}^{(i)} \frac{\eta p(z_t \mid m_{t-1}^{(i)}, x_t^{(i)}) p(x_t^{(i)} \mid x_{t-1}^{(i)}, u_{t-1})}{p(x_t^{(i)} \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, z_t, u_{t-1})} \quad (5)$$

$$\propto w_{t-1}^{(i)} \frac{p(z_t \mid m_{t-1}^{(i)}, x_t^{(i)}) p(x_t^{(i)} \mid x_{t-1}^{(i)}, u_{t-1})}{\frac{p(z_t \mid m_{t-1}^{(i)}, x_t^{(i)}) p(x_t^{(i)} \mid x_{t-1}^{(i)}, u_{t-1})}{p(z_t \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, u_{t-1})}} \quad (6)$$

$$= w_{t-1}^{(i)} \cdot p(z_t \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, u_{t-1}). \quad (7)$$

Unfortunately, the optimal proposal distribution is in general not available in closed form or in a suitable form for efficient sampling. As a result, most efficient mapping techniques use a Gaussian approximation of the optimal proposal. This approximation can easily be computed and allows the robot to sample efficiently.

For each particle i, the parameters $\mu_t^{(i)}$ and $\Sigma_t^{(i)}$ of the Gaussian are determined individually by sampling J test points $\{x_j\}_{j=1}^J$. The test points have to be sampled close to the expected location of the robot. A good guess for the expected location of the robot can be determined by scan-matching, a method to find the pose with a locally optimal match of the current scan with the map constructed so far. *Note that a robust scan-matching method is an important prerequisite for successfully applying the RBPF mapping technique.* Otherwise, the Gaussian approximation that is used as the proposal distribution is not valid. The Gaussian is computed for each sample based on the set $\{x_j\}_{j=1}^J$

$$\mu_t^{(i)} = \frac{1}{\eta^{(i)}} \sum_{j=1}^J x_j \cdot p(z_t \mid m_{t-1}^{(i)}, x_j) p(x_j \mid x_{t-1}^{(i)}, u_{t-1}) \quad (8)$$

$$\Sigma_t^{(i)} = \frac{1}{\eta^{(i)}} \sum_{j=1}^J p(z_t \mid m_{t-1}^{(i)}, x_j) p(x_j \mid x_{t-1}^{(i)}, u_{t-1})$$

$$\cdot (x_j - \mu_t^{(i)})(x_j - \mu_t^{(i)})^T, \quad (9)$$

with the normalization factor

$$\eta^{(i)} = \sum_{j=1}^J p(z_t \mid m_{t-1}^{(i)}, x_j) \cdot p(x_j \mid x_{t-1}^{(i)}, u_{t-1}). \quad (10)$$

With Eq. (8)-(10), we obtain a closed form approximation of the optimal proposal which enables us to efficiently obtain the next generation of particles. Using this proposal distribution, the weights have to be computed as

$$w_t^{(i)} = w_{t-1}^{(i)} \cdot p(z_t \mid m_{t-1}^{(i)}, x_{t-1}^{(i)}, u_{t-1}) \quad (11)$$

$$= w_{t-1}^{(i)} \cdot \int p(z_t \mid m_{t-1}^{(i)}, x') \cdot p(x' \mid x_{t-1}^{(i)}, u_{t-1}) \, dx'$$

$$\simeq w_{t-1}^{(i)} \cdot \eta^{(i)}. \quad (12)$$

As we showed in previous work [6], such an approach is well suited to robustly learn accurate maps of the environment with wheeled robot such as ActivMedia Pioneer robots equipped with SICK laser range finders.

IV. ADAPTATIONS FOR MAPPING WITH A HUMANOID

In this section, we present the modifications to our mapping system so that it can be applied to successfully learn accurate maps with data acquired by a humanoid.

A. Attitude-based Scan Correction

When applying the approach described in the previous section to laser data recorded by a humanoid, the quality of the map will typically be poor. One reason for this is that the humanoid cannot move and at the same time keep the sensor with zero roll and pitch angles. As a result, the object that caused a reflection of the laser beam is observed at different heights in consecutive scans. This makes it nearly impossible to match scans as it is a prerequisite for our approach. Hence, we have to account for the changing attitude. The roll and pitch angle of the laser range finder, however, can be quite accurately estimated using an attitude sensor or a low-cost IMU. As result, we are able to compute the three-dimensional position of the object that caused a reflection. Let ρ_k be the measured range of laser beam k and α_k the corresponding angle. The 3D position of the object is given by

$$\begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \rho_k R(\phi, \theta, \psi) \begin{pmatrix} \cos \alpha_k \\ \sin \alpha_k \\ 0 \end{pmatrix} + \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix}, \quad (13)$$

where R is the 3×3 rotation matrix, $\phi, \theta,$ and ψ refer to the roll, pitch, and yaw angle, and r_x, r_y, r_z to the position of the sensor in the world.

By assuming that the observed objects are walls, cupboards, or similar objects that have the same shape for all z, we can compute a corrected range ρ_k' as

$$\rho_k' = \rho_k \left\| \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} R(\phi, \theta, 0) \begin{pmatrix} \cos \alpha_k \\ \sin \alpha_k \\ 0 \end{pmatrix} \right\|. \quad (14)$$

By replacing ρ_k by ρ_k' for each beam k in a laser range scan, we obtain a corrected range observation that stays constant under changes to the attitude of the sensor. Note that if the roll and pitch angle are too huge, scans might hit the floor. In this case, the measurements are neglected since they do not contain information about the walls.
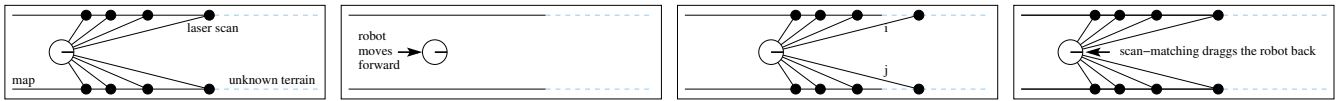
Fig. 2. The problem of scan-matching in a feature-less corridor: (1) The scan is integrated, (2) the robot moves forward, (3) the new scan is obtained, (4) the scan is matched against the map constructed so far and thus the robot is dragged back to the previous location.

B. Scan-Matching with Poor Features

To compute the Gaussian proposal in our mapping system, we perform scan-matching to find the most likely pose by matching the current observation against the map constructed so far

$$x_t^* = \underset{x}{\arg\max}\, p(x \mid m_{t-1}, z_t, x_t'), \qquad (15)$$

where x_t' is the initial guess which is typically computed from odometry. In practice, one applies Bayes' rule and seeks for the pose with the highest observation likelihood $p(z_t \mid m_{t-1}, x)$. Often, a search technique similar to gradient descent is applied. To compute the likelihood of an observation, we use the so called "beam endpoint model". In this model, the likelihood of an individual beam is computed based on the distance between the endpoint of the beam and the closest obstacle from that point.

Using the Hokuyo URG scanner, we have a significantly reduced measurement range compared to traditional range scanners such as a SICK LMS. As a result, the robot can observe only a small area of the environment. Especially if a mobile robot moves along a symmetric structure with poor features, such as a long corridor, scan-matching becomes difficult since ambiguities cannot be resolved. Fig. 2 illustrates a typical situation in which scan-matching in a feature-less corridor fails.

In such a situation, observations are identical or at least very similar for all poses independent of the horizontal position. As a result, the scan-matching procedure reverts the movements of the robot in the horizontal direction. The reason for this is that the obtained scans match perfectly the map at the initial position. Therefore, corridors are often shorter in the maps than in reality. In case an accurate odometry information is available, such situations might be resolved. In general, the shorter the range of the laser, the higher the risk of ambiguities and the lower the quality of the odometry, the worse the initial guess for the matching algorithm. Since this problem occurs comparably often when building maps with a humanoid robot equipped with a Hokuyo URG scanner, we present a way to better deal with such ambiguities when matching scans.

To overcome this problem, we propose to use only a subset of the scan for finding the correspondence but to use the full scan to update the map. Even if this might sound counter-intuitive at first sight, it substantially improves the result of scan-matching in areas with poor features. By neglecting a small fraction of the scan, namely those beams that are close to the maximum usable range of the scanner, the problem of virtually dragging the robot backwards during the pose correction step can be reduced. If, however, in Fig. 2 the beams labeled as i and j are not used for matching but for
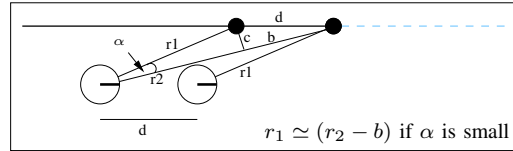

Fig. 3. Geometry used to determine the maximum length of a beam that should be used for scan-matching.

updating the map, the scan-matcher will typically confirm the predicted position and will mainly align the robot in the vertical direction but not in the horizontal one. Note that after this correction is carried out, the full scan is used to update the map.

Based on this example, we can investigate which beams should be neglected in the matching phase. Consider the situation depicted in Fig. 3 in which the robot moves a distance d forward. After moving, the robot should only consider those beams for matching that are likely to hit an obstacle that was visible from the previous location of the robot. To consider only these obstacles, we have to neglect the beams that are longer than the maximum usable range of the scanner minus the distance b. Let α be the angular resolution between two beams. Then, we can compute the length b as

$$b \;=\; \sqrt{d^2 - c^2} \;\;=\; \sqrt{d^2 - \sin^2 \alpha \cdot r_1^2} \le d. \qquad (16)$$

As can be seen, b is bounded by the distance d moved by the robot. As a result, we can improve the scan-matching in feature-less corridors by using only those scans which are shorter than the maximum usable range of the scanner minus the estimated moved distance d of the robot when optimizing Eq. (15).

Note that an alternative strategy that neglects beams that end in an unknown cell in the previous map is not sufficient since this can lead to wrong corrections if appropriate features are visible and the motion of the robot is slightly overestimated.

C. Dealing with Missing Odometry

Today's expensive humanoid robots provide odometry information that allows a robot to incrementally build local maps (compare Gutmann *et al.* [7]). Low cost or self constructed humanoids, however, often do not provide such a reliable estimate about the motion of the robot. Dead reckoning, i.e., the prediction of the robot's pose based on executed motion commands, could be applied, however, it provides only very noisy estimates due to slippage on the ground. Our robot, for example, does not provide any usable information about its relative movement.
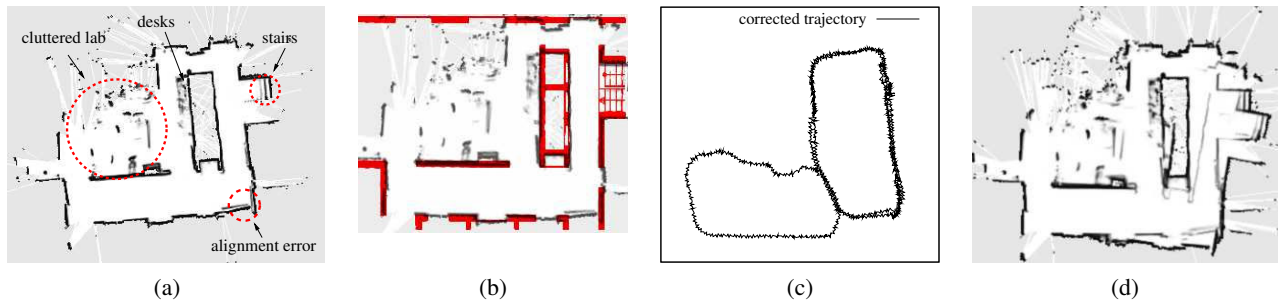
Fig. 4. Image (a) depicts the grid map learned with data acquired by a humanoid. By considering an overlay with the floor-plan (walls in the floor plan are colored red), one can see that the environment is rather accurately mapped (b). Image (c) shows the corrected trajectory of the laser scanner projected on the x/y plane. Image (d) depicts a map learned without the attitude correction. The map is inconsistent and comparably blurred.

Reducing only the range during scan-matching as described in the previous paragraph is not sufficient when no odometry information at all is available. We furthermore need an initial estimate for the motion of the robot to obtain a good pose and thus map estimate. Assuming that the robot starts in a place at which sufficient distinct structure is available for scan-matching, it can quite accurately estimate its motion. Assuming an approximately constant speed of the vehicle, we set

$$u_{t-1} \;\;=\;\; x_{t-1}^* \ominus x_{t-2}^*, \tag{17}$$

where u_{t-1} is the motion estimate that guides the robot from x_{t-1} to x_t, $\ominus$ is the inverse of the motion composition operator, and x^* results from Eq. (15). This estimate leads to a reasonable odometry guess in case the robot moves with constant speed through passages that do not provide distinct features in the sensor data. In all other passages, the scan-matching technique will anyway find an acceptable pose estimate. In practice, Eq. (15) and Eq. (17) are always computed in an alternating way. One starts with scan-matching without odometry, then estimates the odometry, which is in the subsequent step used to initialize the scan-matcher. Even though, the techniques presented in this paper do not describe a new mapping framework, we found that they are relevant to solve the mapping problems with humanoid robots. Furthermore, we believe that they can be easily integrated into other mapping frameworks which apply scan-matching as an intermediate procedure.

V. Experiments

The humanoid robot Robotinho used for the experiments is depicted in Fig. 1. It is self-constructed, around 1 m tall with a total weight of about 5.2 kg, and has 23 degrees of freedom. For our experiments, we equipped it with a Hokuyo URG laser range finder and a XSens IMU (here, only the attitude information is used). The Hokuyo URG is a light-weight scanner with a maximum range of 4.2 m. A measurement range of 4 m, however, can only be obtained with bright and highly reflective obstacles. Dark doors, badly reflecting furniture, or even grayish concrete walls lead to a significantly reduced measurement range of the scanner if the obstacle is not hit perpendicularly. The XSens is used to estimate the attitude (roll and pitch angle) of the chest of the robot. The robot does not possess any odometry sensors.

Fig. 5. Grid map learned from data obtained by a human carrying a laser range finder in the hand. The arrows indicate parts of the environment in which the sensor perceived mainly invalid observations. The estimated trajectory is shown in red.

A. Learning Accurate Grid Maps

The first experiment is designed to show that our approach is well suited to learn accurate grid maps with a humanoid robot. We steered our robot with a joystick through our lab environment. It consists of two corridors which result in two loops the robot traversed, one of them three times. As can be seen in Fig. 4 (a) and (b), our system maps the environment rather accurately. Only one small alignment error occurred at a part of the wall/door. One interesting observation is that the stairs can be identified quite well in the map. The parallel lines are not alignment errors but result from reflection of the individual steps while the robot was walking and thus changing the attitude of the sensor. Fig. 4 (c) shows the estimated trajectory of the robot during that experiment.

Additionally, we disabled the attitude-based scan correction to illustrate its effect. The right image in Fig. 4 depicts the result. The map is more blurred since the laser beams often hit the wrong grid cells. Furthermore, the filter made one wrong pose correction which leads to an inconsistent estimate.

We furthermore performed a second experiment in which a person was holding the laser range scanner and was walking through the environment. As can be seen in Fig. 5, also here we obtain a comparably good map. By looking closely, one can see that the corridors are not perfectly matched and one corridor is slightly too short. This is due to the fact that on one side, the corridor consists of glass panes only and the Hokuyo scanner does not provide any valid data in case a beam hits glass. Therefore, several observations contain no useful information which results in an imperfectly aligned
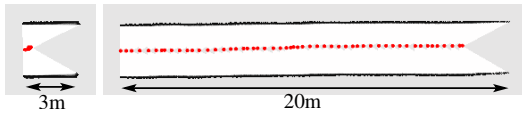
Fig. 6. Scan-matching in a feature-less corridor. Left: Using standard scan-matching, the estimated pose of the robot is always the same. Right: By using our approach, the robot performs much better even if the corridor is still too short (20 m vs. 22.3 m). The estimated trajectory is shown in red.

beam length	forward movement d				
reduction	0.3m	0.4m	0.5m	0.75m	1.0m
0.3m	19.9m	16.3m	14.1m	10.1m	9.5m
0.4m	19.7m	20.0m	17.2m	11.8m	9.6m

Fig. 7. Estimated corridor length for different movements (truth=22.3 m).

map. Nevertheless, the resulting map is sufficient for most tasks such as robot localization or motion planning.

B. Scan-Matching with Poor Features

A further set of experiments investigates the advantages of our scan-matching variant compared to the same approach lacking our technique. By neglecting long beams during the matching phase but integrating them into the map, a substantially better pose estimate can be obtained. Fig. 6 shows the result of the standard scan-matching approach (left image) and our variant (right image). As can be seen, our approach does not provide a perfect motion estimate since the corridor is still shorter than in reality (20 m instead if 22.3 m). As illustrated in Fig. 7, the estimated forward movement (d in Eq. (16)) is a good parameter to quantify the reduction of the maximum valid beam length. If the beam length reduction is chosen too small, the accuracy of the scan matcher drops substantially.

C. Limitations

Given the short range of the sensor, our robot is currently only able to map environments without large free spaces such as hallways. In large rooms such as entrance halls, the proposal cannot be computed accurately - especially without real odometry. In case the robot moves through areas with poor structure, we assumed a constant speed of the vehicle. Furthermore, our system might be less accurate in situations in which the attitude is affected by significant changes while at the same time most of the objects observed by the scanner look different in different heights such as desks for example. In our current configuration, however, the system appears to be comparably robust in practical scenarios.

VI. CONCLUSION

In this paper, we addressed the problem of learning accurate grid maps using laser data acquired by a humanoid. We present techniques to deal with the specific difficulties of typical humanoids such as changing roll and pitch angle of the sensor, missing odometry information, and comparably noisy and short-range sensor data. As a result, we are able to apply a Rao-Blackwellized particle filter to estimate the joint posterior about the trajectory of the robot and the map of the environment. In combination with the adaptations for the mentioned difficulties, this solution to the simultaneous

localization and mapping problem allows a humanoid robot to robustly learn maps. As our experimental results show, the resulting grid maps have a high accuracy, similar to maps built with a wheeled robot. To the best of our knowledge, our system is the first one which is able to build such accurate grid maps containing several loops with a humanoid robot.

REFERENCES

[1] M. Bennewitz, C. Stachniss, W. Burgard, and S. Behnke. Metric localization with scale-invariant visual features using a single perspective camera. In *European Robotics Symposium 2006*, volume 22 of *STAR Springer tracts in advanced robotics*, pages 143–157, 2006.
[2] A. Davison, I. Reid, N. Molton, and O. Stasse. MonoSLAM: Real-time single camera SLAM. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 29(6), 2007.
[3] A. Doucet. On sequential simulation-based methods for bayesian filtering. Technical report, Signal Processing Group, Dept. of Engeneering, University of Cambridge, 1998.
[4] A. Eliazar and R. Parr. DP-SLAM: Fast, robust simultainous localization and mapping without predetermined landmarks. In *Proc. of the Int. Conf. on Artificial Intelligence (IJCAI)*, pages 1135–1142, Acapulco, Mexico, 2003.
[5] P. Elinas, R. Sim, and J. J. Little. σSLAM: Stereo vision SLAM using the rao-blackwellised particle filter and a novel mixture proposal distribution. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, 2006.
[6] G. Grisetti, C. Stachniss, and W. Burgard. Improved techniques for grid mapping with rao-blackwellized particle filters. *IEEE Transactions on Robotics*, 23(1):34–46, 2007.
[7] J.-S. Gutmann, M. Fukuchi, and M. Fujita. A floor and obstacle height map for 3D navigation of a humanoid robot. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, Barcelona, Spain, 2005.
[8] D. Hähnel, W. Burgard, D. Fox, and S. Thrun. An efficient FastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 206–211, 2003.
[9] S. Julier, J. Uhlmann, and H. Durrant-Whyte. A new approach for filtering nonlinear systems. In *Proc. of the American Control Conference*, pages 1628–1632, Seattle, WA, USA, 1995.
[10] G. Lidoris, K. Kühnlenz, D. Wollherr, and M. Buss. Information-based gaze direction planning algorithm for SLAM. In *Proc. of IEEE-RAS Intl. Conf. on Humanoid Robots (Humanoids)*, 2006.
[11] M. Montemerlo and S. Thrun. Simultaneous localization and mapping with unknown data association using FastSLAM. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, pages 1985–1991, Taipei, Taiwan, 2003.
[12] K. Murphy. Bayesian map learning in dynamic environments. In *Proc. of the Conf. on Neural Information Processing Systems (NIPS)*, pages 1015–1021, Denver, CO, USA, 1999.
[13] J. Neira and J.D. Tardós. Data association in stochastic mapping using the joint compatibility test. *IEEE Transactions on Robotics and Automation*, 17(6):890–897, 2001.
[14] R. Ozawa, Y. Takaoka, Y. Kida, K. Nishiwaki, J. Chestnutt, J. Kuffner, S. Kagami, H. Mizoguchi, and H. Inoue. Using visual odometry to create 3d maps for online footstep planning. In *Proc. of IEEE Intl. Conf. on Systems, Man, and Cybernetics*, 2005.
[15] R. Smith, M. Self, and P. Cheeseman. Estimating uncertain spatial realtionships in robotics. In I. Cox and G. Wilfong, editors, *Autonomous Robot Vehicles*, pages 167–193. Springer Verlag, 1990.
[16] S. Thompson, S. Kagami, and K. Nishiwaki. Localisation for autonomous humanoid navigation. In *Proc. of IEEE-RAS Intl. Conf. on Humanoid Robots (Humanoids)*, 2006.
[17] S. Thrun, W. Burgard, and D. Fox. A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping. In *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, San Francisco, CA, 2000.